NPS-AS-94-001

NAVAL POSTGRADUATE SCHOOL

Monterey, California

Alternative Learning Curve Models:
An Analysis of Forecast Error

O. Douglas Moses

January 1994

Approved for public release; distribution unlimited

Prepared for: Naval Center for Cost Analysis
Washington, D.C. 20350-1100

NAVAL POSTGRADUATE SCHOOL
Monterey, California

This report was prepared in conjunction with research conducted for The Naval Center for Cost Analysis and funded by the Naval Postgraduate School.

This report was prepared by:

O. Douglas Moses
Associate Professor
Department of Systems Management

Reviewed by:

for

David R. Whipple, Chairman
Department of Systems Management

Released by:

Paul J. Marto
Dean of Research

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
January 1994

3. REPORT TYPE AND DATES COVERED
Technical Report, January 1994

4. TITLE AND SUBTITLE
Alternative Learning Curve Models:
An Analysis of Forecast Error

5. FUNDING NUMBERS
O&MN, Direct Funding

6. AUTHOR(S)
O. Douglas Moses

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Department of Systems Management
Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-AS-94-001

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
Numerous learning curve models have been offered in the literature and used in practice. This paper selects five learning curve models which differ with respect to the pattern of learning assumed to exist, and investigates the forecast accuracy of the varying models under varying circumstances. The broad objectives are to (1) identify conditions which may affect model accuracy, documenting the manner in which forecast errors for each model depend on those conditions, and (2) suggest which of the five models may be more or less accurate under a given set of conditions. Tests are conducted using annual cost and quantity data from a large sample of major aerospace weapon system programs.

14. SUBJECT TERMS
Learning curve, Cost progress model, Cost estimation, Forecasting accuracy

15. NUMBER OF PAGES
42

16. PRICE CODE
N/A

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std 23908
298-102

ALTERNATIVE LEARNING CURVE MODELS:

AN ANALYSIS OF FORECAST ERROR

O. Douglas Moses
Associate Professor
Department of Systems Management
Naval Postgraduate School
Monterey, CA 93943
408-656-3218

January 1994

ALTERNATIVE LEARNING CURVE MODELS:
AN ANALYSIS OF FORECAST ERROR

ABSTRACT

Numerous learning curve models have been offered in the literature and used in practice. This paper selects five models which differ with respect to the pattern of learning assumed to exist, and investigates the forecast accuracy of the models under varying circumstances. The broad objectives are to (1) identify conditions which may affect model accuracy, documenting the manner in which forecast errors for each model depend on those conditions, and (2) suggest which of the five models may be more or less accurate under a given set of conditions. Particular attention is paid to how model accuracy is affected by one specific condition -- changes in production rate.

ALTERNATIVE LEARNING CURVE MODELS:
AN ANALYSIS OF FORECAST ERROR

INTRODUCTION

Learning curve models have been widely discussed and widely used in practice to estimate costs expected during a repetitive production or acquisition program (Teplitz, 1991; Yelle, 1979). And numerous forms of learning curve models have been suggested and developed (Liao, 1988). This paper evaluates the forecast accuracy of five alternative learning curve models by examining the ability of the five models to estimate future cost during an ongoing production/acquisition program. At the most general level, the objective of the research is to document the relative accuracy of the five alternative models.

Methodology: Broadly, the methodology used to assess the accuracy of the five models is as follows: (a) cost and quantity data from a sample of program was collected, (b) each of the five alternative models was fit to t years of data and then used to estimate (forecast) cost for the next (t+1) year, (c) actual t+1 cost was compared with estimated t+1 cost to measure forecast error, and (d) forecast errors from the different models were observed under various circumstances and statistical tests were employed to draw conclusions concerning the pattern of errors and the conditions that significantly impact the accuracy of each specific model.

Differing Models: The five learning curve models differ from each other in several respects. Specifically, the models differ in

terms of (a) whether they assume significant learning is occurring or not, (b) whether they rely on program-specific learning rates or industry-wide learning rates, (c) whether they rely on "complete" data or only on "recent" data to establish learning rates, and (d) whether they assume a linear or log relationship between cost and cumulative quantity. Identifying such differences between the models permits findings concerning how model characteristics are associated with forecast error. Thus a second objective of the research is to document relationships between model characteristics and model accuracy.

Differing Conditions: A premise of the study is that the accuracy of a model might depend on the conditions in which the model is expected to perform (Conway and Schultz, 1959; Adler and Clark, 1991). The study identifies and creates variables to reflect, seven conditions: (1) the variability in production quantities, (2) the production rate trend, (3) the richness of the data in terms of the number of data points, (4) the degree of learning, (5) the mix of fixed and variable costs in total costs, (6) the period-to-period variability in cost, and (7) the anticipated change in production rate. A third objective of the research is to document if and how model accuracy depends on each of these seven conditions and to suggest which of the five models might perform "best" under which set of conditions.

ALTERNATIVE LEARNING CURVE MODELS

Consider the central purpose of a learning curve model. It is

not really a model that explains cost per se. (It says nothing about the absolute amount of cost.) Rather its purpose is to explain the relationship between costs at different points during a repetitive production/acquisition process. Every learning curve model rests on two assumptions: (1) that future cost depends on past cost, and (2) that future cost differs systematically from past cost as a function of experience gained during the repetitive process. Alternative models differ primarily in terms of what is assumed about the relationship between cost and experience. Five models are offered below, each making different assumptions.

1. Random Walk (RW) Model: The simplest of all, the random walk model assumes that future cost is equal to the most recent past cost:

$$C_{t+1} = C_t \tag{1}$$

where

C = unit cost
t = sequencing subscript

This naive model assumes there is no relationship between cost and experience and serves as a benchmark for assessing the accuracy gained by including additional variables.

2. Traditional Learning Curve (LC) Model: The traditional learning curve is the model most often used for incorporating "experience" into the prediction.

$$C_{t+1} = C_1 \, Q_{t+1}^{b} \tag{2}$$

where

C_1 = theoretical first unit cost
Q = cumulative quantity produced

b = a parameter, the learning curve exponent or slope
C, t = as before

Here the traditional log-linear relationship between C and Q is assumed. C_1 and b are determined for each specific program by fitting the curve to past data. Then C_{t+1} is forecast by plugging in a value for Q_{t+1}. This model assumes a program specific learning rate and uses all available past cost/quantity data to determine that rate.

3. Two Point (TP) Learning Curve Model: Rather than using all past data to estimate a learning rate, this model uses only the two most recent data points. Thus it assumes that only the most recent learning experience is relevant to anticipating the future learning to be expected. Still assuming a log-linear relationship between C and Q, the most recent learning slope is estimated by

$$b = \frac{\log\ (C_t/C_{t-1})}{\log\ (Q_t/Q_{t-1})}$$

Then assuming that future learning will follow the same slope implies

$$C_{t+1} = C_t\ \mathrm{EXP}\ (b\ (\log\ (Q_{t+1}/Q_t))) \qquad (3)$$

where

EXP = exponential function (e to the power in the parentheses).

C, b, Q, t = as before

4. Two Point Linear (LN) Model: Traditionally learning curves have assumed a log-linear relationship between cost and quantity. This model alters that assumption and replaces it with a linear assumption. If cost and quantity are linearly related,

and the slope is estimated using the most recent two points, then the slope would be

$$b' = \frac{C_t - C_{t-1}}{Q_t - Q_{t-1}}$$

and future cost would be forecast by

$$C_{t+1} = C_t + b' (Q_{t+1} - Q_t) \qquad (4)$$

5. Industry (IN) Learning Curve Model: This model assumes there is a standard learning rate within an industry and that that industry rate is more representative of learning that can be expected on a program than is any program-specific rate:

b_I = industry learning rate (the average b of all programs in the sample).

Future cost is then forecast by:

$$C_{t+1} = C_t \ \mathrm{EXP} \ (b_I \ (\log \ (Q_{t+1}/Q_t))).$$

To recap, the assumptions built into the models imply that conceptually the models differ along several dimensions. The five models differ in terms of

a) whether they assume learning is occurring (models 2, 3, 4, 5) or not (model 1).

b) whether they rely on program-specific learning rates (models 2, 3, 4) or an industry-wide rate (model 5).

c) whether learning rates are estimated using "complete" data (models 2, 5) or only "recent" data (models 3, 4).

d) whether learning results in a log-linear relationship between cost and quantity (models 2, 3, 5) or a linear relationship (model 4).

ASSESSING ACCURACY

The objective of the study is to investigate model accuracy under various conditions. The data for the study involved costs and quantities for successive production lots. Accuracy here is defined in terms of the ability of a model to correctly forecast the "next lot average unit cost." Accuracy in such near term cost forecasting is seen as being a relatively minimal requirement expected of a cost progress model. The basic process is quite simple:

(a) Models were fit to a series of cost points to estimate (when necessary) model parameters.

(b) Estimated models were used to forecast future (next period) average unit cost.

(c) Realized actual unit costs were compared to forecasted costs to assess accuracy.

It should be noted here that model accuracy centrally involves the ability to correctly forecast in advance, not the ability to explain a cost series ex post. Two notions of accuracy apply. One is the absolute magnitude of forecast error, regardless of whether the forecast is too high or too low. The second is the direction of the error, whether the model under or over-estimates future cost. Given two concepts, two measures were used:

$$\text{ERROR} = |\text{PUC} - \text{AUC}| \div \text{AUC} \quad (6)$$

$$\text{BIAS} = (\text{PUC} - \text{AUC}) \div \text{AUC} \quad (7)$$

where

PUC = predicted unit cost

AUC = actual unit cost

ERROR is a commonly used accuracy measure, the absolute percentage

error. ERROR can take on only positive values and higher values, of course, signal poorer forecasts. BIAS takes on both positive and negative values. Positive (negative) values signal over (under) prediction of cost.

CONDITIONS AFFECTING MODEL ACCURACY

The general research hypothesis is that the accuracy of models will depend on the circumstances in which they are used. What circumstances might impact accuracy? Prior research (Smunt, 1986; Moses, 1991, 1992) has suggested and discussed variables that might have an effect. Below such variables are listed, with a brief description and comment on how they were operationalized (measured) empirically. Collectively these variables will be referred to as the "condition" variables because they attempt to represent exogenous conditions which may affect model accuracy.

1. Fixed Cost Burden: Total unit cost must consist of both variable costs and a share of the total fixed cost burden associated with capacity. A major role of production rate is determining the volume of output over which fixed capacity costs will be spread. Hence, unit cost will depend on production rate. Learning models ignore this production rate impact on cost, likely causing forecast error. Thus model accuracy may depend on the degree to which total unit cost is made up of fixed costs. The following regression equation was fit to cost series data and the coefficient f used as a measure of fixed cost burden.

$$C_t = v + f \frac{1}{R_t}$$

(This equation is consistent with seeing total unit cost per period (c_t) as the sum of variable cost per unit (v) plus a standard fixed cost per unit (f) adjusted for relative production rate per period (R_t). Higher values of f would be consistent with greater fixed cost burden, i.e., a greater proportion of fixed cost in total cost.)

2. Learning Slope: Past simulation research (Smunt, 1986) shows that the importance of including a learning parameter in a cost model depends, not surprisingly, on the degree of learning that exists in the data. Hence, accuracy across the five models examined may depend on learning rate. Learning slopes were measured by using the b parameter estimated from model 2, transformed to learning rates (e.g., 90%, 80%, etc.). Higher values indicate less learning.

3. Cost Variability: Costs may vary from period to period due to unsystematic random factors. Such random factors influencing cost can be expected to obscure systematic relationships between cost and quantity variables, reducing the chance that a cost model will be estimated correctly and forecast accurately (Smunt, 1986; Moses, 1991). Empirically, Cost Variability was measured by the average period-to-period (lot-to-lot) percentage change in average unit cost. Higher values indicate greater period-to-period variability in unit cost.

4. Quantity Variability: If production rates (lot quantities) are highly unstable across periods, the amount of fixed cost burden assigned to individual units would vary greatly, and

unit cost, would be unstable. Learning rates estimated under such conditions would likely be unreliable, resulting in inaccurate forecasts from learning models. Thus model accuracy may depend on the degree to which production rate/quantity varies. Empirically, Quantity Variability was measured by the average period-to-period (lot-to-lot) percentage change in production quantity. Higher values indicate greater quantity variability.

5. Quantity Trend: When initiating a production/acquisition program for a new item, does production rate (lot quantity) start at a low level and build up slowly to full capacity? Or is full capacity production achieved rapidly? Simulation results (Moses, 1991) have shown that the rate at which lot quantities grow when initiating a program affects cost model accuracy. Does a similar relationship exist when using real data? Empirically, the growth trend in lot quantity was operationalized by dividing first lot quantity by the average lot quantity over the (to date) life of a program. Hence, it is a measure of first lot size as a proportion of average lot size and a crude indicator of the trend in quantity. Lower values indicate greater growth in quantity relative to initial quantity.

6. Plot Points: The number of data points available to estimate the parameters of a model may affect model accuracy. Not surprisingly, simulation results (Moses, 1991) show that when comparing the relative accuracy of models, models with fewer (more) parameters tend to be relatively more accurate when the number of observations is smaller (greater). One question is whether similar

findings will come with real data.

7. Future Production Rate: Once a model is estimated using past data, it is used to forecast future cost. Changes in production rate between the model estimation period and the future should alter future unit cost and hence reduce a model's ability to forecast that future cost accurately. The degree of disadvantage would be expected to depend on how much future production rate differs from the past. Empirically, a variable measuring the change in production rate was constructed by dividing next (future) period's rate by last (most recent) period's rate. (This ratio was then logged to make the distribution symmetrical.) Positive (negative) values indicate increases (decreases) in production quantities.

SAMPLE AND DATA

The accuracy of the cost progress models was investigated using data for a sample of military aircraft and missile systems programs taken from the U. S. Military Aircraft Cost Handbook (DePuy, et. al., 1983) and the U. S. Missile Cost Handbook (Crawford, et. al., 1984). These handbooks contain data for virtually all military aircraft and missile programs from the early 1960s through the early 1980s. Two basic data items were collected from the handbooks for each program: annual lot quantities and average airframe unit costs per lot (in 1981 constant dollars). Programs were deleted from consideration if there were incomplete data or if the programs ran less than five years (a minimum number of data points was needed to fit the models). Based on these

criteria, 46 programs (32 aircraft, 14 missile) were included in the final sample. These programs ranged in length from five years to thirteen years.

The original sample of 46 programs was "expanded" into 121 separate cost series. This was accomplished by dividing each program cost series into separate individual year-to-date cost series. For example, if a particular program had cost data available for six years, say 1970-1975, this single program cost series would be expanded into three separate series as follows:

Cost series #1: 1970-1973 data (used to forecast 1974 cost)

Cost series #2: 1970-1974 data (used to forecast 1975 cost)

Cost series #3: 1970-1975 data (used to forecast 1976 cost)

Thus the initial cost series for each program includes the first four years of data, while subsequent cost series were created by additionally including data from the next year in the cost series. This approach makes maximum use of data and approximates the actual process of a cost estimator who would update a forecast model each period to incorporate the most recent data.

ANALYSIS AND FINDINGS

The basic methodology used to assess cost model accuracy was as follows: Each of the five alternative models was estimated (when necessary) on each of the 121 cost series. Next-period cumulative quantity was input to each model to forecast next-period unit cost. Then next-period forecasted cost and next-period actual cost were compared. Thus the process produced 121 measures of error for each of the five models. The analysis primarily involves

describing and explaining the pattern of errors observed across the different models and across the different circumstances (i.e., across different values of the seven condition variables).

General Error Patterns - Descriptive Statistics:

Table A provides selected descriptive statistics for both ERROR and BIAS for the five models. Some general patterns are evident. On average, the random walk (RW) and industry ning model (IN) produce cost forecasts with the lowest ERROR, with a mean of about 12% and median around 8%. The traditional learning curve (LC) and the two-point learning model (TP) do a little less well and the linear model (LN) has the highest error. Although not shown in the table, the same ordering exists at the 25% and 75% quartiles. This suggests that the relative accuracy of the five models is consistent throughout the distribution of observations, and is not caused by extreme individual observations influencing the average magnitude of error. The same general ordering also exists for the measures of dispersion in errors--standard deviation and range).

Although not universal, there is also a general pattern evident for BIAS. Models 2, 3, 4, and 5 all exhibit negative bias, up to about 6%. Negative bias is a tendency to under-forecast future cost. Models 2, 3, 4, and 5 all assume learning occurs. The negative bias implies that the models anticipate a greater degree of cost reduction than actually occurs, leading to forecasted costs that are lower than those realized.

Table A
Error Statistics for Alternative Learning Curve Models

	MODELS				
Statistic	1) RW	2) LC	3) TP	4) LN	5) IN
Mean-absolute error	.125	.169	.161	.219	.121
Median- absolute error	.074	.124	.109	.140	.084
Stnd. Dev.- absolute error	.129	.153	.145	.217	.122
SIQR[1] -absolute error	.126	.169	.174	.191	.132
Mean-bias	.049	-.033	.003	-.088	-.011
Median-bias	.016	-.061	-.034	-.057	-.034

1. SIQR= Semi-interquartile range: (75th quantile - 25th quantile)

Errors and Model Characteristics

As indicated earlier, the five models differ along several dimensions. Some broad observations about the relationship between model characteristics and the magnitude of forecast error is possible.

First, the RW model, assuming no learning, outperformed (lowest error) the other four models. This is somewhat surprising, given that the sample, aerospace programs, is one where systematic learning is conventionally assumed to occur and thus one where models explicitly incorporating learning would be expected to have an advantage. On average, learning (cost reduction) does occur in the sample (this is evident from the fact that the RW model, ignoring learning, systematically overestimates cost, a positive BIAS), but the fact that the RW model "misses" this learning is less of a detriment to accuracy than the generally greater unreliability of the other four models.

Second, of the models incorporating learning, the industry model (IN) outperforms the three models (LC, TP, LN) which rely on program-specific estimates of learning. This has a somewhat interesting implication: It suggests that if an analyst wishes to project the degree of learning to be expected in the future on an existing program, the past learning experience on that program provides a poorer indication that does the "average" learning experience within the industry.

Third, when constructing and using a program-specific learning model, it is not obvious that all of the data (the full program

cost history) should be used to estimate a learning rate. Note that the two-point (TP) model performs marginally better than the traditional learning curve (LC) model. This suggests that in forecasting near-term future cost reduction, the learning experienced during the most recent past may be more relevant that the learning experienced over a program's full history.

Fourth, if program-specific learning is to be modeled, the conventional assumption of a log-linear relationship between cost and quantity is superior to the alternative linear assumption. This follows from noting the poor performance of the LN model, the highest error overall. Why this is so can be seen by looking at the BIAS measures. All of the learning models have a tendency to under-estimate future cost. Log-linear models assume cost will decline with increasing quantity, but at a decreasing rate; while a linear model assumes cost will decline at a constant rate. This linear assumption simply compounds the negative bias existing for all the learning models, leading to even greater under-estimation of future cost and higher error.

Relationship Between Accuracy and Conditions:

Is the accuracy of the models dependent on the circumstances in which they are used? Do models perform well in some circumstances, less well in others? To get a first-cut answer to these questions, two tests of the relationship between ERROR (from each of the five models separately) and the condition variables were conducted:

1. Pairwise Correlations: This is a univariate test of association, where measurement errors in other variables do not intrude.

2. Multiple Regression of ERROR on the Condition Variables together: This is a test of association for each variable while controlling for the others.

Correlations, regression coefficients and t values from these two approaches are provided in Table B. Several observations concerning Table B follow.

First, where results are strong (significant at a higher level of probability) in one of the two tests, they tend to be corroborated in the other test. So there is at least some convergence across the two tests.

Second, for three of the seven conditions (Quantity Variability, Quantity Trend and Plot Points) there are no significant results and thus no indication that model accuracy depends on these factors. This is of interest simply because all of the factors in this study have been shown to impact accuracy in at least one of the simulation studies cited previously.

Third, significant results are found for the other four condition variables, and these results are not limited to single models. The manner in which these conditions affect the accuracy of the individual models tends to be fairly consistent (although the degree and significance of the relationship differs from model to model.) What follows is a look at the impact of the conditions. The approach used was to partition the sample into three subsamples

depending on whether the values for a condition variable were low (bottom quartile), medium (middle 50%), or high (top quartile) and then, for each model, observe and plot average values for ERROR for these three subsamples. This approach is followed below for variables found significant in the Table B tests.

Error Analysis by Condition

1. Burden: Consider first the results for Burden. All correlations and regression coefficients are positive (although significance is not strong). This general result is as hypothesized and is plotted in Figure A. As factory burden increases, as the proportion of fixed cost in unit cost increases, learning models become less accurate. Because learning models do not incorporate the impact on unit cost of spreading period fixed costs over varying output quantities, forecast errors are expected. And the magnitude of the errors are directly associated with the amount of fixed cost burden.

2. Learning Slope: The Table B regression results indicate that ERROR is positively associated with Learning Slopes. What is not apparent from this positive regression coefficient is that the relationship is not monotonic. Observation of average ERROR by quartile (not shown) indicates that for the four learning models (2, 3, 4, and 5), forecast errors are moderate for the lowest quartile, smaller for middle range values, and largest for the top quartile, i.e., a V-shaped pattern. In short, for the four models incorporating learning, ERROR is higher when estimated learning rates are in either the bottom or top quartiles. A fuller story

Table B
Test of Relationship Between Learning Curve Model Errors and Explanatory Conditions

Conditions	Test Statistics	1) RW	2) LC	3) TP	4) LN	5) IN
Burden:	Corr.	.13	.19*	.17	.20*	.20*
	Reg. Coef.	.05	.01	.05	.06	.08
	Reg. t	1.11	.20	1.06	.78	1.76
Learning Slope:	Corr.	.12	.01	.23*	.09	.17
	Reg. Coef.	.32	.48	.51	.41	.40
	Reg. t	2.20*	2.89**	3.24**	1.67	2.91**
Cost Variability	Corr.	.09	.35***	.18*	.24**	.14
	Reg. Coef.	.06	.35	.21	.30	.10
	Reg. t	.76	3.58***	2.31*	2.11*	1.24
Quantity Variability	Corr.	-.10	.03	-.04	-.01	-.07
	Reg. Coef.	-.07	-.03	-.03	-.05	-.03
	Reg. t	-1.31	-.54	-.50	-.63	-.55
Quantity Trend	Corr.	-.05	.09	-.19*	-.12	-.10
	Reg. Coef.	.04	.06	.01	-.00	.02
	Reg. t	1.31	1.88	.24	-.05	.91
Plot Points	Corr.	.06	.01	-.00	-.07	.07
	Reg. Coef.	.00	.00	.00	-.00	.01
	Reg. t	.64	.25	.38	-.35	.93
Future Production Rate	Corr.	.21*	.09	.15	.19*	.11
	Reg. Coef.	.04	.01	.02	.04	.02
	Reg. t	2.50*	.68	1.08	1.65	1.08

* Significant at .05
** Significant at .01
*** Significant at .001

PLOT OF MODEL FORECAST ERRORS BY LEVEL OF BURDEN

LEGEND:
RANDOM WALK MODEL (MODEL 1) = SERIES 1
LEARNING CURVE MODEL (MODEL 2) = SERIES 2
TWO-POINT MODEL (MODEL 3) = SERIES 3
LINEAR MODEL (MODEL 4) = SERIES 4
INDUSTRY MODEL (MODEL 5) = SERIES 5

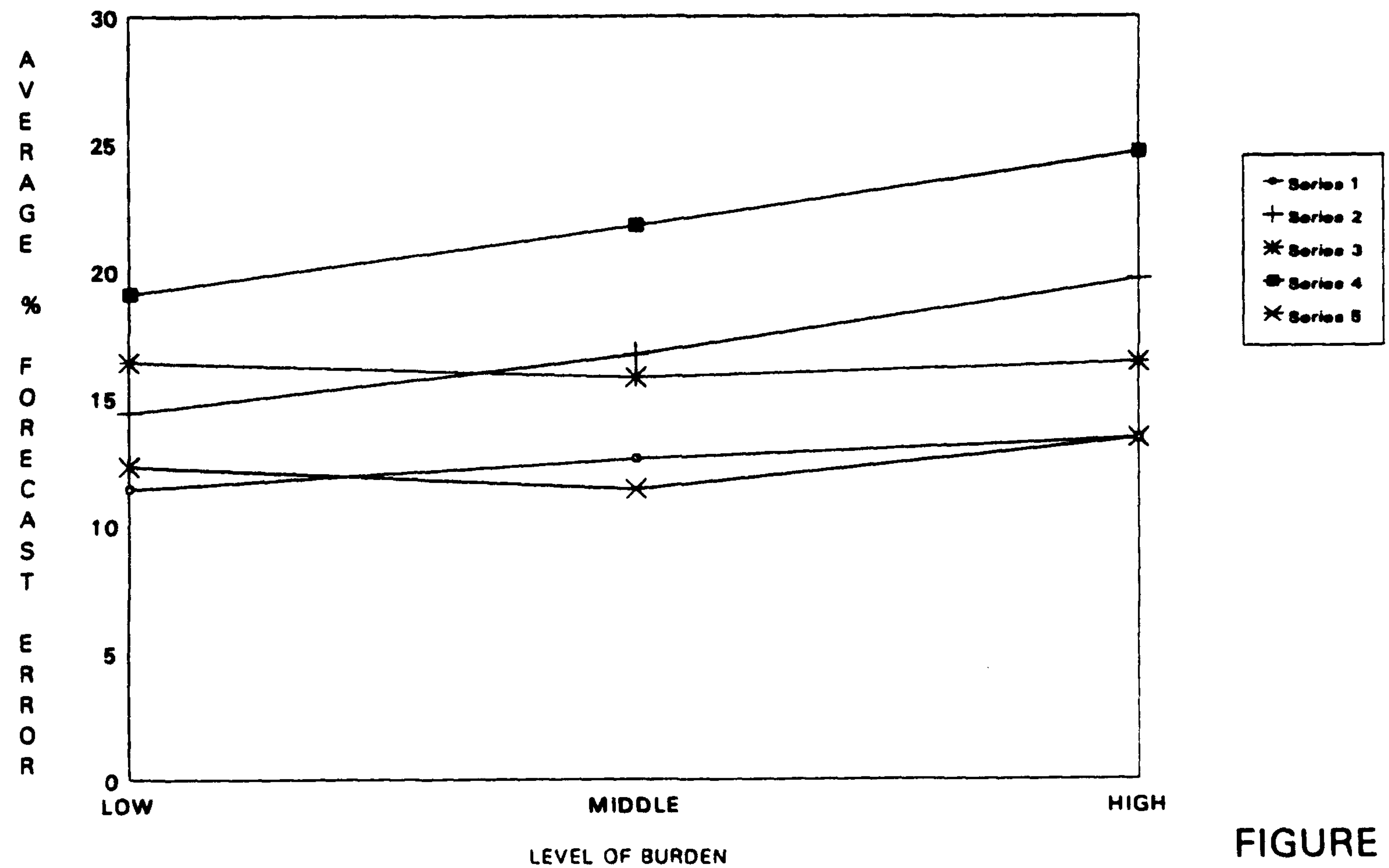

FIGURE A

comes from observing BIAS rather than ERROR. A plot of BIAS by quartiles is shown in Figure B. When much learning appears to be occurring, the learning model under-estimates future cost. When little learning appears to be occurring, the learning models over-estimate future cost. What seems to be happening is a "regression to the mean" effect. A high (low) rate of past cost reduction causes the model to forecast a high (low) rate of future cost reduction and, in each case, the high (low) rate regresses to a more average rate, causing consistent over-or under-estimation of future cost.

3. Cost Variability: The Table B results show a generally positive relationship between ERROR and Cost Variability. Figure C shows that this is caused primarily by a deterioration in model accuracy when past variability in the program cost series has been "high" (the top quartile subsample). This finding is consistent with past simulation results suggesting that learning models try to explain all variability in cost through the estimation of the single learning parameter and, when there is considerable period-to-period "noise" in the cost series, end up erroneously "interpreting" that noise in the estimated learning rate.

What this suggests is that the pattern of unit cost experienced in the past during a program can indicate something about the ability of a learning model to forecast future cost for the program. High variability in past unit cost signals high unreliability in learning curve model forecasts.

The Impact of Future Production Rate

Of the seven condition variables, Future Production Rate is special for two reasons. First, conceptually it is distinct. The other six variables describe conditions existing during the periods over which the models are estimated -- i.e., the past. In contrast, Future Production Rate describes a condition (the level of production) expected to exist during the period for which cost is being forecast. Second, how models perform in situations where production rates are changing is of particular importance for today's cost analyst, facing cost forecasting problems in an environment of rapid industrial change, such as production rate cutbacks in the defense industry.

The table B results concerning the relationship between ERROR and Future Production Rate are not strong, but this is perhaps misleading. Correlations and regressions test for linear relationships and prior research suggests that the relationship may be non-linear. Consider Figure D, plotting mean ERROR versus Future Production Rate. A clear V-shaped pattern exists, with model ERRORs larger for both the top and bottom quartiles of Future Production Rate.

What does the V-shaped pattern mean? Simply put, if production rate in the period for which cost is being forecast diverges much from the recent past, either up or down, the accuracy of all five of the models deteriorates. This is not a surprising finding. All models in the study fail to incorporate any variable to reflect the impact of changing production rate on unit cost.

PLOT OF MODEL FORECAST BIAS BY LEVEL OF LEARNING SLOPE

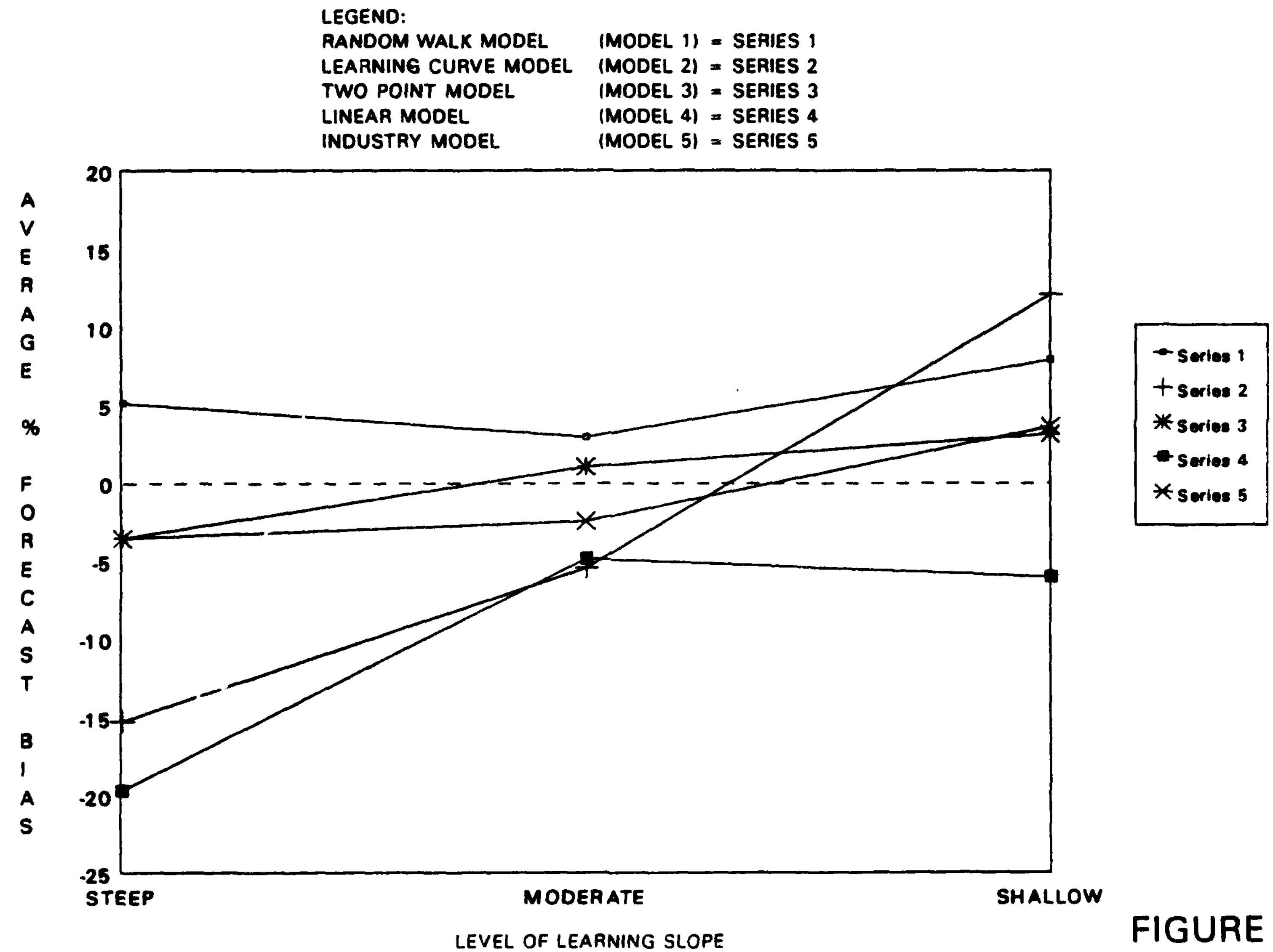

FIGURE B

PLOT OF MODEL FORECAST ERRORS BY LEVEL OF COST VARIABILITY

LEGEND:
RANDOM WALK MODEL (MODEL 1) = SERIES 1
LEARNING CURVE MODEL (MODEL 2) = SERIES 2
TWO POINT MODEL (MODEL 3) = SERIES 3
LINEAR MODEL (MODEL 4) = SERIES 4
INDUSTRY MODEL (MODEL 5) = SERIES 5

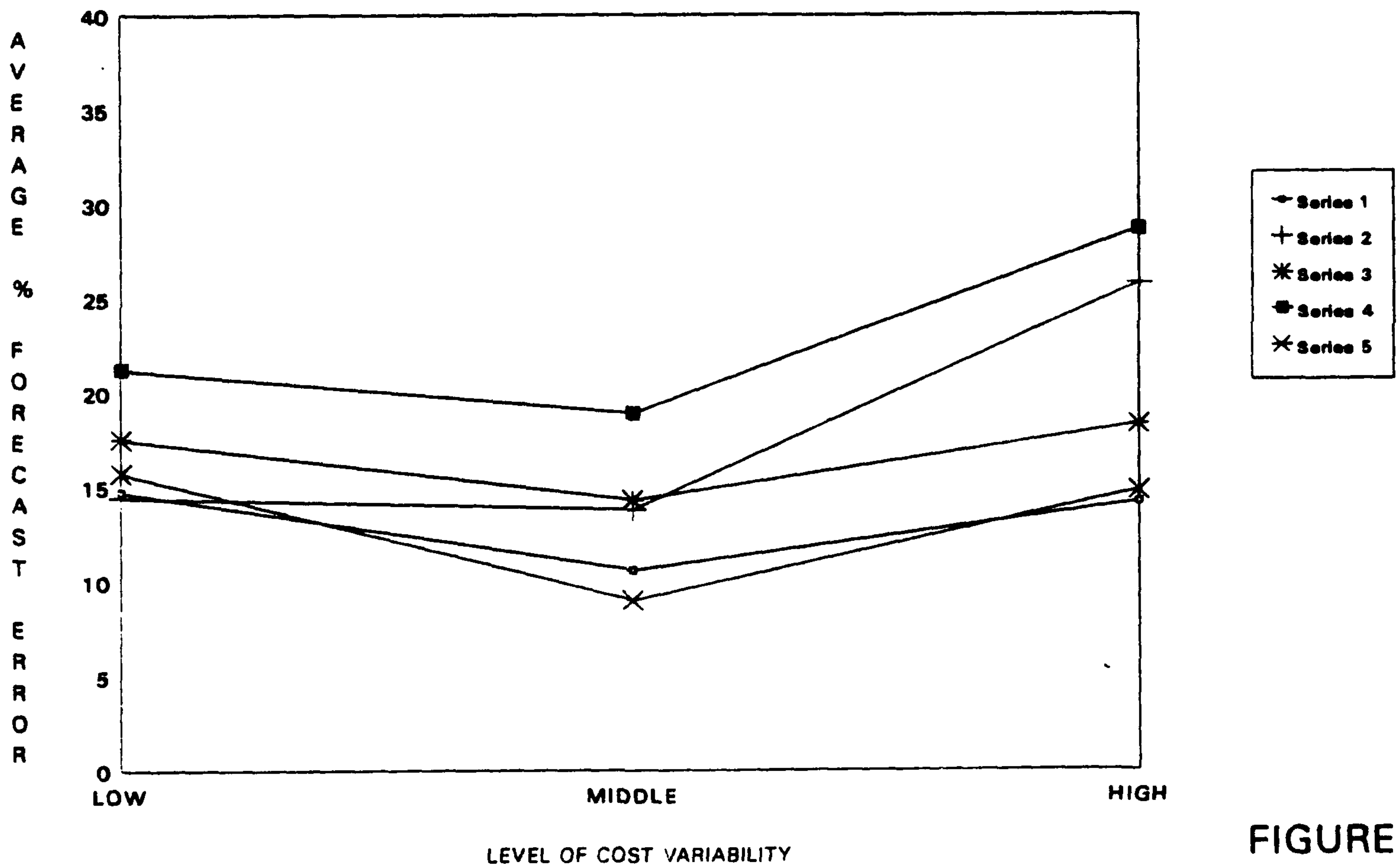

FIGURE C

PLOT OF MODEL FORECAST ERRORS BY LEVEL OF FUTURE PRODUCTION RATE

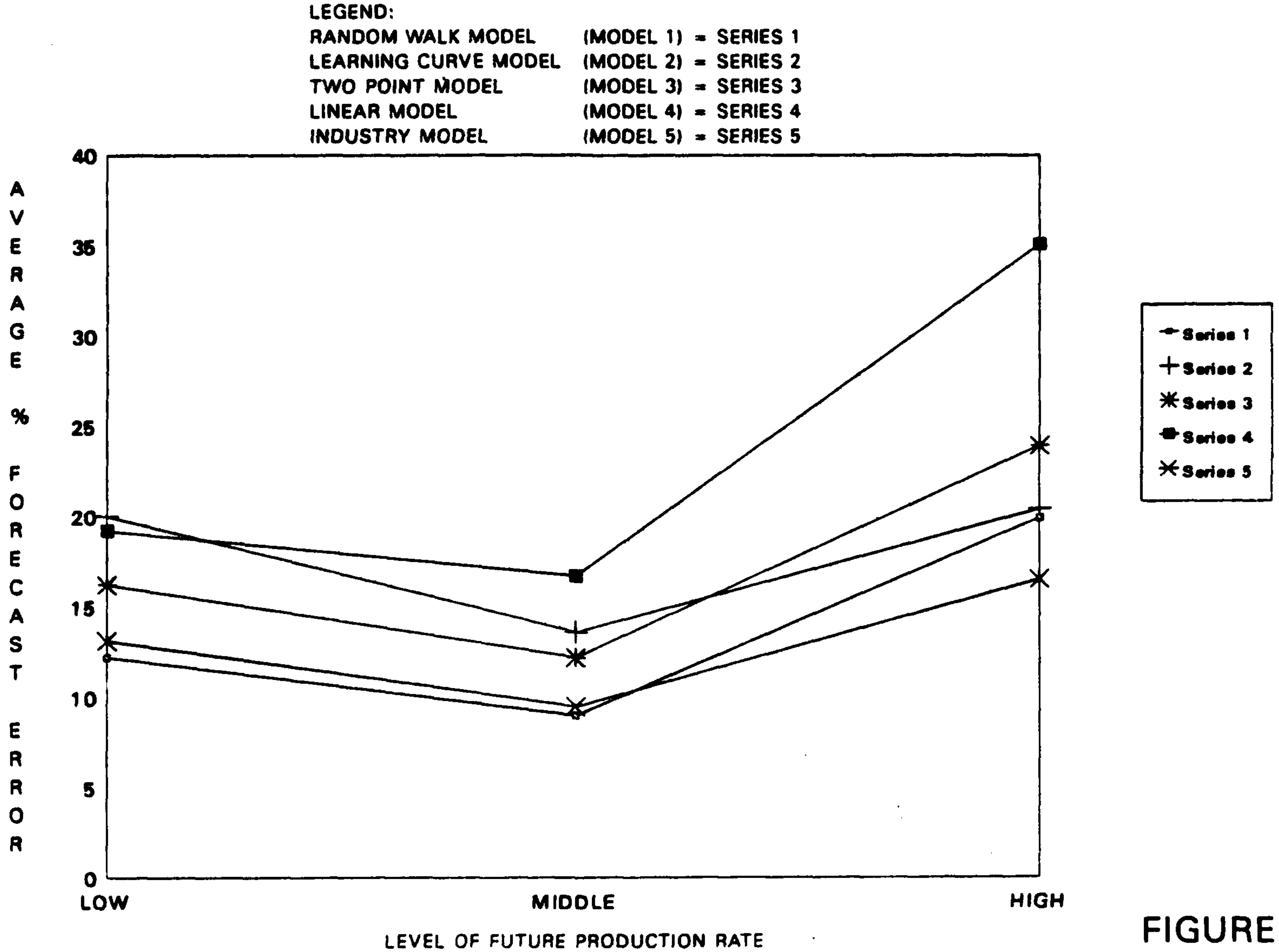

FIGURE D

Given that all the models mis-forecast cost when future production rate changes, a related question is: In what direction? This can be answered by observing values for BIAS, which are plotted in Figure E. The patterns in Figure E is of interest: All models both under-estimate cost (negative BIAS) when future production rate falls and over-estimate cost when future production rate rises. This is not surprising. Falling rate should increase actual unit cost, because fixed capacity costs are spread over less output. The learning models "miss" this effect and thus consistently under-estimate unit cost. The opposite effect occurs when production rate increases, leading to over-estimates of unit cost.

Comparisons of Model Accuracy

Given that the accuracy of the five models depends on the conditions under which they are used, an inevitable question arises: Which model appears to perform "best" under which conditions? Table C ranks the models by median ERROR, both overall (full sample) and by subsamples partitioned on values of the seven condition variables. Several observations seem noteworthy from these comparisons.

First is the consistent domination of the RW model, ranking most accurate overall and in a majority of the subsamples. The primary place where the RW model performs less well is in the subset where Future Production Rate is "up" relative to the past. This is plausible. The RW model has a small bias toward over-

estimation of future cost. When future production rate increase relative to the past, actual realized unit cost will decline (due to spreading fixed costs over increased output volume), magnifying the bias and hence forecast error.

Next, is the "second place" showing for the IN model. It is second most accurate overall and tends to be the model that outperforms the RW when the RW is not most accurate. In fact, the IN model is worse than second best in only one of the subsamples. It appears that the overall superiority of the RW and IN models is not due to superior accuracy under just some conditions; rather, that superiority holds across all variations in the conditions tested.

Third is the tendency for the models that required estimation of a program-specific learning rate (the LC, TP and LN models) to perform less well. Again this finding tends to hold across all the subsamples. Consider the LN model, for example, which has the highest error overall, and performs no better than fourth best out of five in any of the subsamples.

CONCLUSIONS AND FINAL COMMENTS

The objective of this paper has been to document the accuracy of five learning curve models under varying conditions, using cost data from real world programs. Accuracy was evaluated in terms of ability to forecast next-period unit cost. Data consisted of annual lot costs from 46 military aerospace programs, arranged so that models were used to forecast 121 next-period costs. The five

PLOT OF MODEL FORECAST BIAS BY LEVEL OF FUTURE PRODUCTION RATE

LEGEND:
RANDOM WALK MODEL (MODEL 1) = SERIES 1
LEARNING CURVE MODEL (MODEL 2) = SERIES 2
TWO-POINT MODEL (MODEL 3) = SERIES 3
LINEAR MODEL (MODEL 4) = SERIES 4
INDUSTRY MODEL (MODEL 5) = SERIES 5

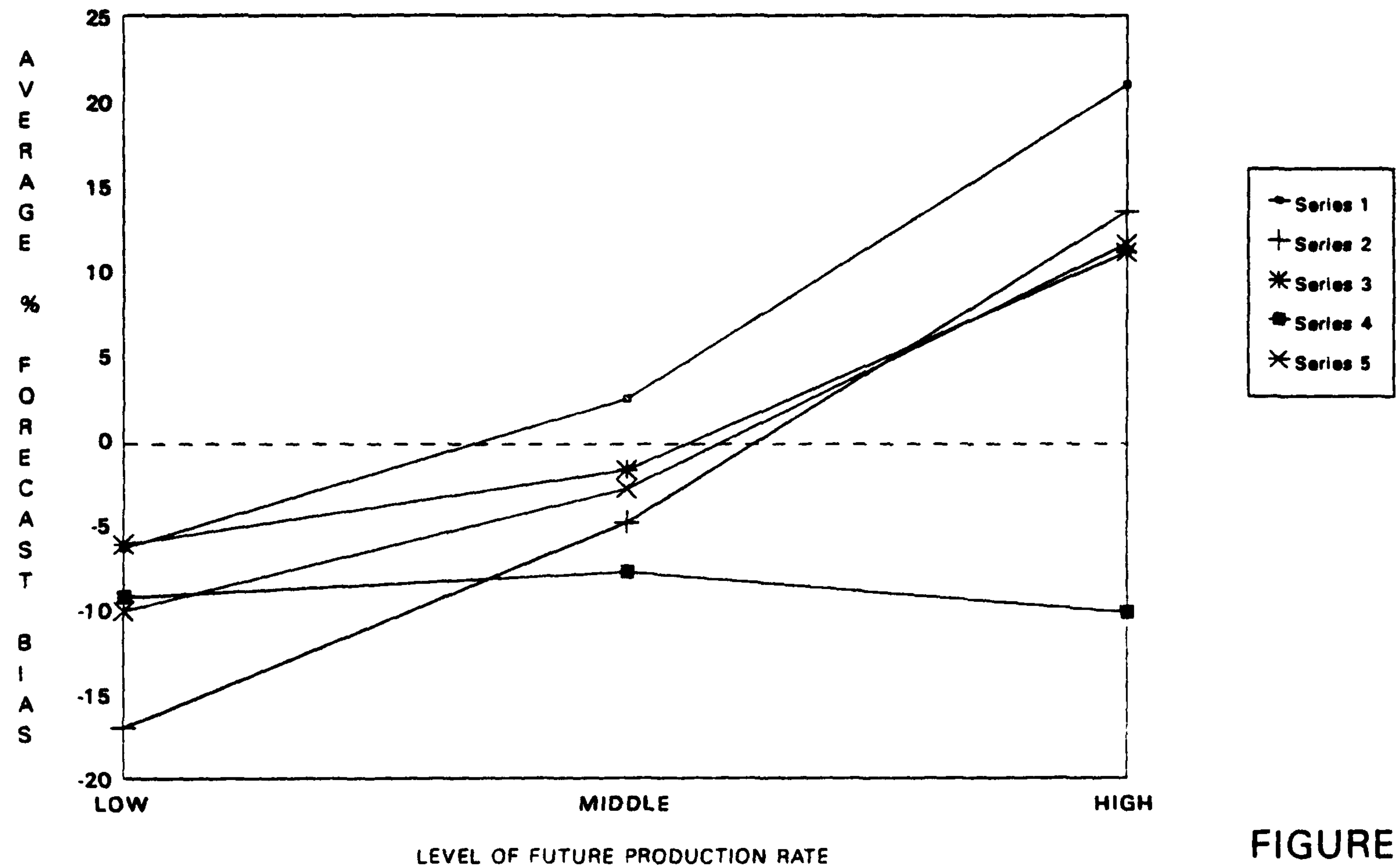

FIGURE E

Table C
Ranking of Alternative Learning Curve
Models in Terms of Median Error
(Most accurate= 1, least= 5)

Conditions:	1) RW	2) LC	3) TP	4) LN	5) IN
Overall	1	4	3	5	2
Burden					
Low	1	3	5	4	2
Moderate	2	5	3	4	1
High	2	3	4	5	1
Learning Slope					
Steep	1	4	3	5	2
Moderate	1	4	3	5	2
Slight	1	3	5	4	2
Cost Variability					
Little	1	2	4	5	3
Moderate	2	4	3	5	1
Great	1	5	3	4	2
Quantity Variability					
Little	1	4	3	5	2
Moderate	3	5	2	4	1
Great	1	4	3	5	2
Quantity Trend					
Little Growth	1	3	4	5	2
Moderate Growth	2	3	4	5	1
High Growth	1	5	3	4	2
Plot Points					
Few	1	3	4	5	2
More	1	5	3	4	2
Many	2	4	3	5	1
Future Prod. Rate					
Down	1	5	3	4	2
Little ch.	1	5	3	4	2
Up	3	2	4	5	1

models forecasted future cost using some combination of variables reflecting (a) past costs, and (b) "experience", although how that experience was modeled differed across the models. Specific findings and error patterns have been presented; broader conclusions follow:

1. The accuracy of all the models (tested) does depend on the circumstances or conditions in which they are used. Those conditions can be identified in advance. Thus a cost estimator using a particular model may be able to assess the risk of forecast error depending on the conditions.

2. Which conditions affect accuracy, and by how much, varies somewhat from model to model. But the results suggest that the amount of fixed cost burden, the degree of apparent learning, the degree of past variability in period-to-period cost, and the nature and degree of change in the future production rate provide information that can inform a cost estimator about the risk of forecast error from using a particular model.

3. It is not obvious that program-specific learning models improve forecasting. Quite the contrary for the sample here; forecast accuracy was best for a random walk or industry learning model.

4. Although a relatively large sample of aerospace programs was included, all of the findings and conclusions should be tempered by the acknowledgement that they came from tests on one set of data -- cost data that was at a high level of aggregation (annual lot costs) and reasonably lean (the maximum data points for

fitting a model was 13). Results would likely be most generalizable to similar cost forecasting situations. On the other hand, many of the error patterns observed in this study have also been observed in previous studies evaluating models on simulated data, so it is unlikely that the error patterns observed can be discounted as simply sample-specific. Perhaps some of the findings may be viewed as tentative -- as hypotheses to be additionally supported (or contradicted) by future research. Given the findings of this study, one direction such research might take would be to start with the following question: Under what circumstances can program-specific learning models outperform a simple random walk or an industry learning model?

References

Adler, P. and K. Clark (1991), "Behind the Learning Curve: A Sketch of the Learning Process," Management Science, Vol. 37, No. 3, March, pp. 267-281.

Conway, R. and A. Schultz (1959), "The Manufacturing Progress Function," Journal of Industrial Engineering, 10, pp. 39-53.

Crawford, D. et. al. (1984), U.S. Missile Cost Handbook, No. TR-8203-3, Management Consulting and Research, Inc., Falls, Church, VA 16 January 1984.

DePuy, W., et. al. (1983), U.S. Military Aircraft Cost Handbook, No. TR-8203-1, Management Consulting and Research, Inc., Falls Church, VA, 1 March 1983.

Liao, S. (1988), "The Learning Curve: Wright's Model vs. Crawford's Model," Issues in Accounting Education, Vol. 3, No. 2, pp. 302-315.

Moses, O. (1991), "Learning Curve and Rate Adjustment Models: Comparative Prediction Accuracy Under Varying Conditions," in R. Kankey and J. Robbins, editors, Cost Analysis and Estimating: Shifting U.S. Priorities, Springer-Verlag, New York, 1991, pp. 65-101.

Moses, O. (1992), "Learning Curve and Rate Adjustment Models: An Investigation of Bias," in T. Gulledge, et al., editors, Cost Analysis and Estimating: Balancing Technology Advances and Declining Budgets, Springer-Verlag, New York, 1992, pp. 3-38.

Smunt, T. (1986), "A Comparison of Learning Curve Analysis and Moving Average Ratio Analysis for Detailed Operational Planning," Decision Sciences, Vol. 17, No. 4, Fall, pp. 475-494.

Teplitz, C. (1991), The Learning Curve Deskbook, Quorum Books, New York.

Yelle, L. (1979), The Learning Curve: Historical Review and Comprehensive Survey," Decisions Sciences, Vol. 10, No. 2, April, pp. 302-328.

Distribution List

Agency	No. of copies
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Dudley Knox Library, Code 52 Naval Postgraduate School Monterey, CA 93943	2
Office of Research Administration Code 012 Naval Postgraduate School Monterey, CA 93940	1
Library, Center for Naval Analyses 4401 Ford Avenue Alexandria, VA 22302-0268	2
Department of Systems Management Library Code SM Naval Postgraduate School Monterey, CA 93943	2
Captain Richard L. Coleman Director, Naval Center for Cost Analysis Room 4A538, The Pentagon Washington, DC 20350-1100	4
Professor O. Douglas Moses Code AS/Mo Department of Systems Management Naval Postgraduate School Monterey, CA 93943	20